Direct Inferences in a Connectionist Knowledge Structure

S. C. Hollbach

Abstract

A model of human cognition is proposed in which all concept properties are context dependent. Concepts are comprised of multiple facets, each motivated by a different functional property. A connectionist implementation is presented in which conceptual modification yields the 'direct inferences' implicit in the structure of a knowledge base.

Introduction

When a noun is modified by a descriptive adjective, the result is often a significant modification of the original concept denoted by the noun. For example, while a peach is soft, juicy and tangy, a green peach is hard, dry and bitter. Clearly, the adjective 'green' when applied to a peach conveys more than merely the colour.

This paper advances a computational model of conceptual modification that captures the 'direct inferences' arising from property correlations. The domain of this investigation is concrete nouns and their attendant descriptive adjectives. Each noun denotes a concept, where a concept is represented by a structured collection of properties and values, indexed by function. For example, an apple viewed as food brings different properties to mind than an apple viewed as a projectile. Ultimately it is the goals and plans of the agent that determine how an object is thought of: a hungry agent thinks of apples differently than an angry one. Thus the functional properties of an object provide the context for interpretation, in that they select only the currently relevant facet of the complete description.

We have built a connectionist implementation of the functional context-sensitive model of category representation. The system runs on a Sun Workstation as an application of the Rochester Connectionist Simulator [Goddard 1987], the results of which are visible in iconic form thanks to the Graphics Interface [Lynne 1987]. The system uses an extensive knowledge base of categories and their interrelations to draw direct inferences about modified categories, answer queries about object properties, and model property dominance effects [Whitney 1986, Tabossi 1986].

The Structure of Knowledge

This investigation focusses on the mental representation of physical objects. The building block of these mental representations are *categories*, classifications of physical objects sharing one or more common properties. A *property* is a set of descriptors applicable to a physical object, where the elements of the set are *property values*. We classify properties into three groups: perceptual, constitutive and functional. Perceptual properties pertain to the five senses, functional properties relate to an object's usefulness by humans, and constitutive properties are in some sense the definitional properties of a category, often expressed in terms of genetics, compositional makeup and so on. Functional properties play a special role in category representation, supplying as they do the various perspectives from which the category can be viewed. For example, the *is edible* property provides the focus of relevance for the *tangy* and *has seeds* properties of apples.

REPORT DOCUMENTATION PAGE

Form Approved
OPM No. 0704-0188

1. AGENCY USE ONLY *(Leave Blank)*

2. REPORT DATE
1989

3. REPORT TYPE AND DATES COVERED
Unknown

4. TITLE AND SUBTITLE
Direct Inferences in a Connectionist Knowledge Structure

5. FUNDING NUMBERS
DAAB10-86-C-0567

6. AUTHOR(S)
S.C. Hollbach

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of Computer Science
University of Rochester
Rochester, NY 14627

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army CECOM Signals Warfare Directorate
Vint Hill Farms Staion
Warrenton, VA 22186-5100

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
92-TRF-0029

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Statement A, Approved for public release, distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

A model of human cognition is proposed in which all concept properties are context dependent. Concpts are comprised of multiple facets, each motivated by a different functional property. A connectionist implementation is presented in which conceptual modification yields the 'direct interfernce' implicit in the structure of a knowledge base.

14. SUBJECT TERMS
Direct Interfences, Artificial Intelligence, Data Fusion, Connectionist Knowledge Structure, Graphics Interface

15. NUMBER OF PAGES
7

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

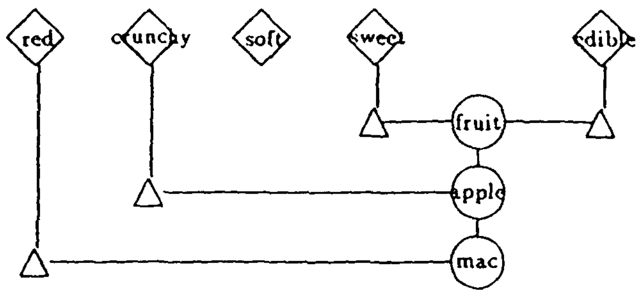

Figure 1: Circles represent categories, diamonds property values and triangles binder nodes associating categories with their attendant values. Lines do not represent direct links, but rather indirect connections mediated by subnet structures of varying complexity.

Categories, in addition to having a complex internal structure, are related to one another in a hierarchical subsumption taxonomy. (A familiar example of such structuring is ontological knowledge, the ordering of natural kinds according to common biological characteristics.) The links in the taxonomy represent subsumption relations between categories. Thus to some degree a lower level category participates in the higher level category. The form this participation takes differs depending on whether one is looking up or down the taxonomy. All the properties and values possessed by the higher level category are also possessed by the lower level one. And for each property or value possessed by a lower level category it is true of the higher level category that there exists an element of that set having that property or value. For example, since all things have colour as a property, all apples must also have a colour. Furthermore there exists a red (or green or yellow) apple, by virtue of the colour values associated with the various apple varieties.

The connectionist implementation of this cognitive model follows the 'localist' paradigm of Feldman and Ballard [1982]. Each category is represented by a single exemplar object. Each object, property and value is represented by a distinct (named) network node. All relations between these nodes are captured in separate subnets of regular structure, allowing the network to be compiled from a series of high level input language statements. Concepts are represented as patterns of activity over all the nodes in the network. 'Thoughts' are formed in the network by keying in activation on a noun and (optionally) adjectives, and allowing the simulation to run a few steps to permit these activations to propagate fully. Activation flows out from the noun denoting the category to all relevant properties and values. Relevance is determined by context, or more specifically, by the currently active functional property of the category. Each category has associated with it a default context or facet; for example, the default view of 'apple' is *edible*. So when the noun is activated in isolation, the system responds by selectively activating its defining properties and values with respect to its default context. For example, Figure 1 depicts the graphics display of certain key elements of the network after keying in activation on the 'apple' node and allowing the simulation to run a few steps to stability. The selective effects of context mean that only a subset of all possible property values of the concept are active at one time, although a given property value can participate in any number of facets.

Direct Inferences

A major feature of the connectionist knowledge base is its dynamic nature. Rather than having the knowledge encoded in a purely passive (declarative) format requiring a distinct reasoning component to interpret and apply it, the knowledge encoded in a connectionist network incorporates several simple forms of inferencing directly in the structure of the network. These direct inferences, that is, inferences not requiring an interpreter but contained entirely within the terminological component,

can be either *mediated* or *immediate*. Immediate inferences are drawn about object properties at the level of the object itself, while mediated inferences involve property inheritance. Mediated inferences can be drawn either from more general knowledge, or, if the information is not available at a higher level, a weaker answer can be derived from more specific knowledge.

Both forms of direct inference, mediated and immediate, arise from this fundamental mode of operation of the network, as demonstrated in Figure 1. One of the many immediate inferences drawn about apples is the fact that they are crunchy; one of the mediated inferences is the fact that they are edible. A further inference, as to the existence of red apples, is obtained by ranging down the hierarchy, rather than up as is customary.

Property Queries

Property queries take the general form "does (modified) category x have property value y?". Phrased more naturally, this becomes "are y's x?" or "do y's have x's?". Of course, given that each category is represented by a single exemplar, a more accurate portrayal of the query forms would be to say "is a y x?" or "does a y have x?", for example, "is a black bird large?" or "does a red apple have seeds?". Each property value is represented by two nodes, one corresponding to the adjective as a category modifier, the other corresponding to a query on that property value. So to pose a query to the system, the user activates the adjective and noun forming the target category, thus invoking the fundamental mode of operation of the network, namely, the drawing of direct inferences. The queried property value is then keyed in on the query-specific twin of the property value node.

There are five possible responses to a query: a 'yes' or 'no' in context, a 'yes' or 'no' out of context, and 'category error'. An 'yes' in context occurs when the property value is an element of the set characteristic of the current facet. A 'no' in context occurs when a modifier negates the queried value, as in "is a green apple red?". A 'yes' out of context is reported when a shift of context is required to answer in the affirmative, as in "are sweet apples easily thrown?". An answer of 'no' out of context results when the property value is not associated with the category in any context, although other values of that property are, as in the query "are small apples purple?". Answers out of context take a little longer than answers in context, since context shifting takes time. A category error occurs when the property associated with the value is not a property of the category ("are ideas purple?"). Category errors are interesting not only for their possible role in cognitive development [Keil 1979], but also for the fact that when they occur in conversation it is generally to signal a metaphor. We are currently in the process of extending the model to account for metaphoric interpretation of such category errors.

Once the query has been keyed in, the simulation is run to a point where either an answer of 'yes' or 'no' in context or a report of a category error would be detected by the system. If neither condition exists, alternate facets are explored in parallel until either an answer of 'yes' out of context is reported or the possibilities are exhausted, resulting in a 'no' (out of context).

Property Dominance Effects

In addition to participating in a subsumption taxonomy, the mental representation of an object has a complex internal structure. The design of this internal structure is based on the premise that all object properties are context dependent. This idea arose from the debate between Whitney [1986] and Tabossi [1986] over the problem of lexical access modelling, or the question of whether the meaning of an ambiguous word is selected at the lexical access stage or interpreted later. Whitney presents related results concerning the semantic access of unambiguous words in support of the multiple access model (akin to the delayed interpretation model for ambiguous words). In Whitney's work, all the properties germane to a concept (as denoted by a concrete noun) are accessed or primed

in parallel, by mention of the noun, regardless of any bias built into the sentential context. The effect of the bias, to promote some properties to prominence and inhibit others, is only visible several hundred milliseconds after inital mention of the nouns, and must thus be occurring at a later (post-lexical) processing stage. Tabossi, on the other hand, contends that the stimuli used in Whitney's work are too neutral with respect to the target concept to induce any significant bias, and presents results to support the competing notion of selective access, or lexical level biasing and inhibition of properties.

The question of whether or not lexical access is context sensitive is still an open one. There is agreement in the literature, however, on the fact that certain concept properties are correlated, both positively and negatively [Malt and Smith 1984], and that context is used to decide which of a number of competing property associations or coalitions should be permitted to dominate [Cohen and Murphy 1984]. As Tabossi points out, ice is both hard and cold, yet a sentence like "The bartender served the drinks with ice" that primes the property value 'cold' also inhibits the value 'hard' and vice versa, while a neutral sentence neither primes nor inhibits either property.

These results are consistent with the model advanced in this work, in that competing contexts are mutually exclusive, neutral contexts are unrelated, and *reinforcing* contexts are mutually excitatory. Dominance effects are modelled by asymmetric link weights between the two competing facets, permitting a high dominance property (*e.g.*, ice temperature) to exhibit stronger effects than a low dominance one (*e.g.*, ice hardness).

A fundamental assumption underlying this work is that categories, as mental constructs of active agents, are inseparably linked with the agent's planning goals. These goals or situational contexts are so influential on the mental structure of categories that a category is meaningless when out of context. Since categories (indeed, all ideas) are by definition meaningful, they must carry with them a *default context* to supply meaning in the absence of other information. Very often, particularly for physical objects, this default context is simply visual recognition. When the word 'apple' is spoken, a mental image is conjured up of the visual appearance of an apple. If the agent is hungry at the time, the apple's taste might spring to mind. If the agent is angry, its properties as a handy projectile might leap into significance. And so on. Thus while a category appears stable to the agent, since the same basic set of properties and values are being drawn on at all times, the structure is actually dynamic, shaped by context.

Contexts interact amongst themselves in different ways than do categories. Where categories can combine with each other in an arbitrarily complex fashion, contexts afford less latitude. A context is nothing more than a particular way of looking at a category. Two perspectives on a category are either the same or different. So contexts can be either mutually supportive or mutually inhibitory, depending on whether they are compatible or not. For example, if the agent's goal is to eat an apple, he must first locate one visually. Thus the *edible* and *visual-id* contexts are compatible. On the other hand, if the agent decides to throw the apple at a passing car, any thought of eating it will be suppressed, as throwing and eating are mutually inhibitory contexts.[1]

The general characteristics of the model derive partly from the psychological model described above and partly from properties of connectionist models. Categories are represented by the various patterns of activity over the set of properties and values associated with that category. Each distinctive pattern is characteristic of a different context or goal. There is a default context associated with each category. If in the process of specifying the structure of the knowledge base the user fails to name a default context for a category, the 'visual-identification' context is used, since this is generally appropriate for the chosen domain of physical objects.

Property values can be either *neutral* or *biasing*. A neutral value displays no strong correlation with one context over any other, while a biasing value is characteristic of only one context. For example, the modifier 'red' is neutral in the phrase 'red pillow' but biasing in the phrase 'red

[1] This version of events is admittedly simplistic, but it suffices for the problem at hand.

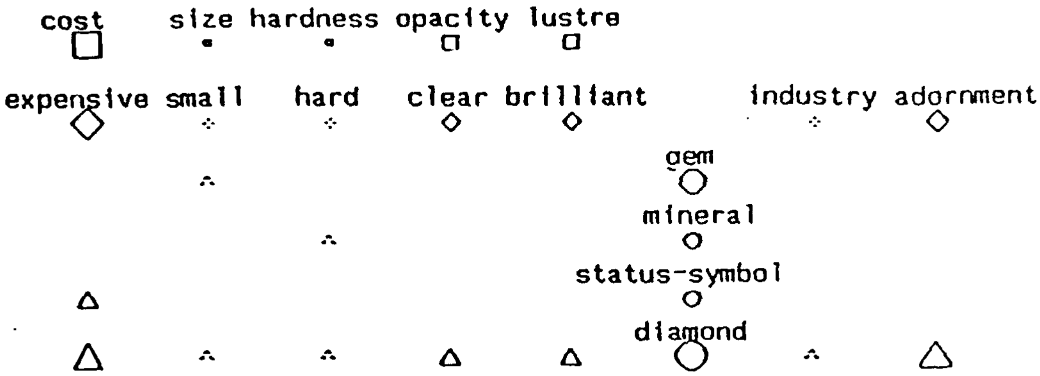

Figure 2: Graphics Interface depiction of changing contexts: initial state.

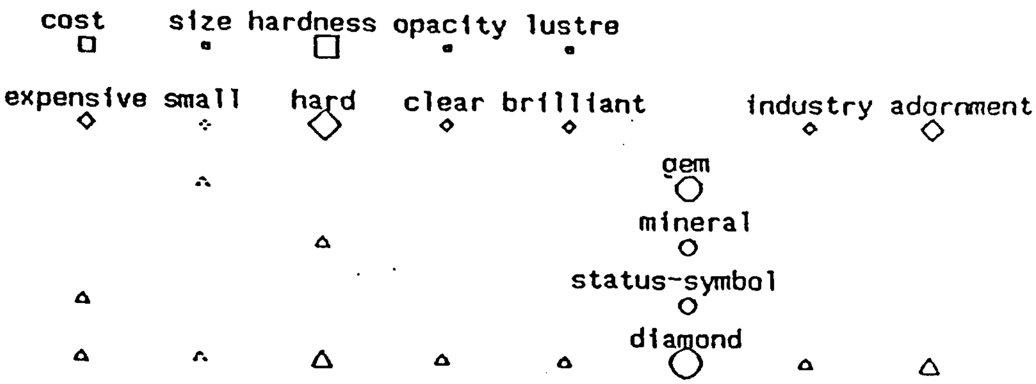

Figure 3: Intermediate state.

rose', raising as it does visions of romance and long-stemmed floral offerings. Thus context can be established implicitly by mentioning a biasing property value; it can also be established explicitly, by turning on the context node by hand. Not all property values are biasing, or, more accurately, not many property values are sufficiently biasing to override the currently active or the default context in favor of another. The values biased toward the context can be guessed at with greater confidence than the more neutral ones, although there is a slight bias built into neutral values. In fact, the neutral/biasing distinction is not a very good one, as it represents the attempt to quantify a gradual change. A more accurate characterization would be to speak of strong, moderate, weak and negligible biases.

Results of running the simulation are shown in Figures 2, 3 and 4. Shown are the iconic representations of individual network nodes. Activation was keyed on the phrase 'expensive diamond' and the simulation allowed to run to stability. Figure 2 shows this initial state of the network. Figure 3 shows the network in an intermediate state shortly after changing contexts from the *adornment* aspect of diamonds to the *industrial* aspect, achieved by shifting the external activation of 'expensive' over to 'hard'. That is, the system is being forced to consider the phrase 'hard diamond' after being primed with the phrase 'expensive diamond'. As Tabossi's property dominance studies predict, there is a significant latency period between presentation of the stimulus and recognition of its appropriateness, as shown in Figure 4. The details of the implementation, including a description of a connectionist interpreter that translates statements in a high level language into a structured network, are given in [Hollbach 1988].

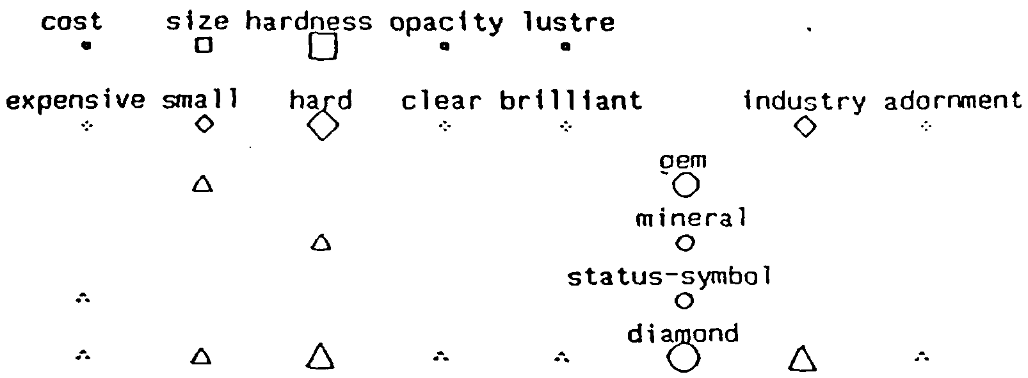

Figure 4: Final state of the network, after a shift in context from 'expensive' to 'hard'.

Summary and Conclusions

This research advances a context sensitive model for conceptual modification and uses it to capture not only property dominance effects but also direct inferences, both immediate and mediated. Concepts are denoted by concrete nouns, optionally modified by one or more descriptive adjectives. The agent's current plans and goals supply the relevance criteria for focussing on a coherent subset of a given object's disparate properties and values. Goals are represented simply by a concept's functional properties. A characteristic use of an object will dominate uncommon ones, as will all perceptual and constitutive properties associated with that functional property, leading to property dominance effects. The interproperty associations yield immediate inferences, and property inheritance gives rise to mediated inferences. The connectionist implementation of this model operates as a question-answering system, permitting the user to pose queries about the various properties of a concept.

Acknowledgments

This work was supported by U.S. Army contract no. DAAB10-87-K-022. Thanks to Jerry Feldman, and to Gary Dell and Josh Tenenberg for invaluable comments.

References

[Cohen and Murphy 1984] Benjamin Cohen and Gregory L. Murphy, "Models of Concepts", *Cognitive Science*, 8:27–58, 1984.

[Feldman and Ballard 1982] Jerome A. Feldman and Dana H. Ballard, "Connectionist Models and Their Properties", *Cognitive Science*, 6:205–254, 1982.

[Goddard 1987] Nigel Goddard, "The Rochester Connectionist Simulator User Manual", Technical Report, Computer Science Department, University of Rochester, April 1987.

[Hollbach 1988] Susan C. Hollbach, Automatic compilation of a connectionist knowledge base, 1988, submitted to AAAI'88.

[Keil 1979] Frank C. Keil, *Semantic and Conceptual Development An Ontological Perspective*, Harvard University Press, Cambridge, Mass., 1979.

[Lynne 1987] Kenton Lynne, "Graphics Interface to the Rochester Connectionist Simulator", Technical Report, Computer Science Department, University of Rochester, April 1987.

[Malt and Smith 1984] Barbara C. Malt and Edward E. Smith, "Correlated Properties in Natural Categories", *Journal of Verbal Learning and Verbal Behaviour*, 23:250–269, 1984.

[Tabossi 1986] Patrizia Tabossi, Effects of context on the immediate interpretation of unambiguous nouns, 1986, Universita di Bologna, Italy. Unpublished.

[Whitney 1986] Paul Whitney, "Processing category terms in context: Instantiations as inferences", *Memory & Cognition*, 14(1):39–48, 1986.